



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/333,121	06/14/1999	MICHAEL E. PALMER	50269-026	3744

7590 11/01/2002

BRAIN D HICKMAN
HICKMAN PALERMO TROUNG & BECKER LLP
1600 WILLOW STREET
SAN JOSE, CA 95125-5106

EXAMINER

BIENEMAN, CHARLES A

ART UNIT PAPER NUMBER

2176

DATE MAILED: 11/01/2002

7

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/333,121

Applicant(s)

PALMER ET AL.

Examiner

Charles A. Bieneman

Art Unit

2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 14 June 1999.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-37 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-37 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 14 June 1999 is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on _____ is: a) ☐ approved b) ☐ disapproved by the Examiner.
- If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
- a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☒ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____.
- 4) ☐ Interview Summary (PTO-413) Paper No(s) _____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: _____.

Art Unit: 2176

DETAILED ACTION

1. This action is responsive to the following communication: original application filed on June 14, 1999.
2. Claims 1-37 are pending. Claims 1, 34, and 35 are independent claims.

Drawings

3. This application has been filed with informal drawings which are acceptable for examination purposes only. Formal drawings will be required when the application is allowed.

Specification

4. The disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code. Applicant is required to delete the embedded hyperlink and/or other form of browser-executable code. See MPEP § 608.01.
5. The disclosure is objected to because of the following informalities:

The word "that" on page 12, line 21 should be deleted.

In listing sources of similarity information the specification lists "CLICK THROUGH SIMILARITY" as number "4" on page 11 and "TITLE SIMILARITY" as number "6" on page 14 without listing any source as number "5".

Appropriate correction is required.

6. The abstract of the disclosure is objected to because it exceeds 150 words in length. Correction is required. See MPEP § 608.01(b).

Claim Rejections - 35 USC § 112

7. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it

Art Unit: 2176

pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

8. **Claims 19-25 and 28-32** are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

Regarding **claims 19-25**, the specification does not describe what in the similarity matrix constitutes structural information, and one skilled in the art would not know what structural information comprises.

Regarding **claims 28-32**, the specification does not describe the operation or construction of a second matrix, and one skilled in the art would not be enabled to practice the invention of claim 28 insofar as it contains this limitation.

The dependent claims not mentioned above are rejected for fully incorporating the deficiencies of their base claims.

9. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

10. **Claims 19-25 and 28-32** are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Regarding **claims 19-25**, Claim 19 recites the limitation "structural information" in lines 1-2, and claims 20 and 21 recite the same limitation in line 1. There is insufficient antecedent basis for this limitation in the claim. The specification does not describe structural information,

Art Unit: 2176

nor is it clear from the face of the claim what subset of information in the similarity matrix comprises structural information.

The dependent claims not mentioned above are rejected for fully incorporating the deficiencies of their base claims.

Regarding **claims 28-32**, claim 28 recites the limitation "the second matrix" in line 3. There is insufficient antecedent basis for this limitation in the claim. Neither claim 1, from which claim 28 depends, nor the specification describe a second matrix.

The dependent claims not mentioned above are rejected for fully incorporating the deficiencies of their base claims.

Claim Rejections - 35 USC § 102

11. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

(e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

The changes made to 35 U.S.C. 102(e) by the American Inventors Protection Act of 1999 (AIPA) do not apply to the examination of this application as the application being examined was not (1) filed on or after November 29, 2000, or (2) voluntarily published under 35 U.S.C. 122(b). Therefore, this application is examined under 35 U.S.C. 102(e) prior to the amendment by the AIPA (pre-AIPA 35 U.S.C. 102(e)).

Art Unit: 2176

12. **Claims 1-12, 19, 26, and 33-37** are rejected under 35 U.S.C. 102(a) and (e) as being anticipated by U.S. Patent Number 5,835,905 to Pirolli et al. (hereinafter "Pirolli"), issued November 10, 1998, filed April 9, 1997.

Regarding **independent claim 1**, Pirolli discloses a method of categorizing a plurality of new documents into a set of categories (Pirolli, col. 3, lines 19-27), each of the categories containing a set of training set documents inasmuch as Pirolli teaches selecting web pages for activation, the selected web pages being equivalent to training set documents. (Pirolli, col. 5, lines 35-42 – "For relevancy predictions, one or more Web pages for spreading activation are selected, step 105. The selected Web pages may be based on the category that it is in. Alternatively, if a user is currently browsing the pages in the web locality, the selected page may be the one currently being browsed. In any event, activation is spread using the selected page as a focal point to generate a list of relevant pages, step 106.")

Further Pirolli discloses that its method uses a matrix representing document similarity that is derived by combining two or more measures of document similarity inasmuch as Pirolli states inasmuch as Pirolli states that "[a]n activation network can be represented as a graph defined by matrix R , where each off-diagonal element $R_{i,j}$ contains the strength of association between nodes i and j , and the diagonal contains zeros." (Pirolli, col. 11, lines 36-39; *see also* col. 8, lines 8-13 – "In order to perform categorizations each Web page at the Web locality is represented by a vector of features constructed from the above topology, meta-information, usage statistics and paths, and text similarities. These Web page vectors are collected into a matrix. Such a matrix is illustrated in FIG. 5.")

Regarding **dependent claim 2**, Pirolli discloses the measures of document similarity including hyperlink similarity inasmuch as Pirolli teaches that the matrix contains “inlinks, the number of hyperlinks that point to the item from the web locality (column 504) [and] outlinks, the number of hyperlinks the item contains that point to other items in the web locality (column 505).” (Pirolli, col.8, lines 19-23; Fig. 5.)

Regarding **dependent claim 3**, Pirolli discloses documents considered similar to each other when there is a link from one to the other, or when the two documents link to, or are linked to by, a set of other associated documents inasmuch as Pirolli teaches that the matrix contains “inlinks, the number of hyperlinks that point to the item from the web locality (column 504) [and] outlinks, the number of hyperlinks the item contains that point to other items in the web locality (column 505).” (Pirolli, col.8, lines 19-23; Fig. 5.)

Regarding **dependent claim 4**, Pirolli discloses certain hyperlinks having greater or lesser similarity weight than other hyperlinks based on other features of the links or their source or destination documents inasmuch as Pirolli teaches “an approach based on weighted linear equations that define the rules for predicting degree of category membership for each page at a web locality. That is, equations are of the form

$$(1) \quad c_i = w_1 v_1 + w_2 v_2 + \dots + w_n v_n$$

for all pages i in a Web locality, where the v_j are the measured features of each Web page, and the w_j are weights.” (Pirolli, col. 8, lines 41-48.)

Regarding **dependent claim 5**, Pirolli discloses the measures of document similarity including a similarity of text of the documents. (Pirolli, Fig. 5.)

Regarding **dependent claim 6**, Pirolli discloses two documents being considered similar based on a comparison of word vectors derived from the text of each of the two documents. (Pirolli, col. 7, lines 57-65 – “The token information is then used to create a document vector, where each component of the vector represents a word, step 403. Entries in the vector for a document indicate the presence or frequency of a word in the document. The steps 401-403 are repeated for each Web page in the Web locality. For each pair of pages, the dot product of these vectors is computed, step 404. The dot product . . . produces a similarity measure.”)

Regarding **dependent claim 7**, Pirolli discloses text similarity determined in part based upon weight values assigned to words of the text, and wherein certain words have greater or lesser weight than other words inasmuch as Pirolli teaches that “entries in the vector for a document indicate the presence or frequency of a word in the document.” (Pirolli, col. 7, lines 59-61.)

Regarding **dependent claim 8**, Pirolli discloses the measures of document similarity including user click-through similarity inasmuch as Pirolli teaches that one “kind of graph[], or network[], . . . used to represent strength of associations among Web pages [is] the usage paths, or flow of users through the locality.” (Pirolli, col. 10, lines 59-60; Fig. 11.)

Regarding **dependent claim 9**, Pirolli discloses documents associated by frequency of clicks inasmuch as Pirolli states that “[r]eferring now to FIG. 13, for the matrix representation of usage path networks, an entry of an integer strength, $s \geq 0$, in column i row j , indicates the number of users that traversed from page i to page j .” (Pirolli, col. 11, lines 30-34. *See also* Pirolli, col. 7, lines 15-18 – “From the set of paths, a vector that contains each page's frequency

Art Unit: 2176

of requests is generated (i.e. a frequency vector), step 304, along with a path matrix containing the number of traversals from one page to another, step 305.”)

Regarding **dependent claim 10**, Pirolli discloses deriving measures of document similarity from patterns detected in user viewing of the documents inasmuch as Pirolli teaches use of “raw data [that] may be obtained from usage records or access logs of the web locality and by direct traversal of the Web pages in the Web locality” (Pirolli, col. 4, lines 57-60), and further states that “[t]he raw data is comprised of topology information, page meta-information, page frequency path information and text similarity information. . . . Usage frequency and path information indicate how many times a Web page has been accessed and how many times a traversal was made from one Web page to another.” (Pirolli, col. 5, lines 2-10.)

Regarding **dependent claim 11**, as discussed above regarding dependent claim 10, Pirolli discloses user viewing information monitored by a web caching system and stored in a log.

Regarding **dependent claim 12**, Pirolli discloses documents being considered similar based on frequency of viewing inasmuch as Pirolli states that “[r]eferring now to FIG. 13, for the matrix representation of usage path networks, an entry of an integer strength, $s \geq 0$, in column i row j , indicates the number of users that traversed from page i to page j .” (Pirolli, col. 11, lines 30-34. *See also* Pirolli, col. 7, lines 15-18 – “From the set of paths, a vector that contains each page's frequency of requests is generated (i.e. a frequency vector), step 304, along with a path matrix containing the number of traversals from one page to another, step 305.”)

Regarding **dependent claim 19**, insofar as that claim can be understood, Pirolli discloses extracting structural information from the similarity matrix to obtain new documents supported by the set of training documents for each category inasmuch as Pirolli teaches extracting

Art Unit: 2176

information from matrix structures and using a spreading activation technique to “define the degree of predicted relevance of Web pages to the starting set of focus Web pages.” (Pirolli, col. 10, lines 8-35.)

Regarding **dependent claim 26**, Pirolli discloses categories coming from a manually derived taxonomy inasmuch as Pirolli states that “for the classification of Web pages in the web locality, classification characteristics are provided, step 103. The classification characteristics are predetermined “rules” which are applied to the feature vectors of a page to determine the category of the page. For example, it may be desirable to have a classification of web pages as index types (contain primarily links to other pages) or content types (contain primarily information).” (Pirolli, col. 5, lines 12-19.)

Regarding **dependent claim 33**, Pirolli discloses identifying a category of a classification taxonomy of the hypertext system in which a first electronic document is presently classified inasmuch as Pirolli states “for relevancy predictions, one or more Web pages for spreading activation are selected, step 105. The selected Web pages may be based on the category that it is in.” (Pirolli, col. 5, lines 34-36.)

Further, Pirolli discloses storing information that classifies the second electronic document into the category if the second electronic document is found to be highly Similar inasmuch as Pirolli states that “activation is spread using the selected page as a focal point to generate a list of relevant pages, step 106.” (Pirolli, col. 5, lines 40-42.)

Regarding **independent claim 34**, Pirolli discloses a computer-readable medium carrying one or more sequences of instructions. (Pirolli, col. 13, lines 24-27.)

Art Unit: 2176

Further, Pirolli discloses the instructions causing processors to execute the step of categorizing a plurality of new documents into a set of categories (Pirolli, col. 3, lines 19-27), each of the categories containing a set of training set documents inasmuch as Pirolli teaches selecting web pages for activation, the selected web pages being equivalent to training set documents. (Pirolli, col. 5, lines 35-42 – “For relevancy predictions, one or more Web pages for spreading activation are selected, step 105. The selected Web pages may be based on the category that it is in. . . . In any event, activation is spread using the selected page as a focal point to generate a list of relevant pages, step 106.”)

Further Pirolli discloses the instructions causing processors to execute the step of using a matrix representing document similarity that is derived by combining two or more measures of document similarity. (Pirolli, col. 8, lines 8-13 – “In order to perform categorizations each Web page at the Web locality is represented by a vector of features constructed from the above topology, meta-information, usage statistics and paths, and text similarities. These Web page vectors are collected into a matrix. Such a matrix is illustrated in FIG. 5.”)

Regarding **independent claim 35**, Pirolli discloses a method of categorizing a plurality of new electronic documents for use in a hypertext search system. (Pirolli, col. 1, line 65 – col. 2, line 2; note also that claims 7-13 recite a method.)

Further, Pirolli discloses creating a storing a set of categories for the documents. (Pirolli, col. 9, lines 6-42.)

Further, Pirolli discloses creating and storing a matrix in which rows and columns identify documents and in which each element of the matrix stores a value that represents a similarity among a pair of documents associated with a row and column that intersect at the

Art Unit: 2176

element inasmuch as Pirolli states that “[a]n activation network can be represented as a graph defined by matrix R , where each off-diagonal element R_{ij} contains the strength of association between nodes i and j , and the diagonal contains zeros.” (Pirolli, col. 11, lines 36-39.)

Further, Pirolli discloses deriving each matrix value by combining two or more measures of similarity that are obtained by analysis of the documents inasmuch as Pirolli discloses that the strength of association contained in the matrix R are such measures of similarity by stating that

Three basic kinds of raw data are extracted from a Web locality:

Topology and meta-information, which are the hyperlink structure among Web pages at a Web locality and various features of the pages, such as file size and URL.

Usage frequency and usage paths, which indicate how many times a Web page has been accessed and how many times a traversal was made from one Web page to another.

Text similarity among all text Web pages at a Web locality

As described mentioned above with respect to FIG. 1, the raw data is used to construct two types of representations:

Feature-vector representations of each Web page that represent the value of each page on each dimension and which are used in the categorization process

Graph representations of the strength of association of Web pages to one another, which are used in the spreading activation. The graphs are represented using matrix formats.

(Pirolli, col. 5, line 55 – col. 6, line 7.)

Regarding **dependent claim 36**, Pirolli discloses creating and storing a graph of links for each measure of document similarity inasmuch as Pirolli states that “three kind of graphs, or networks, are used to represent strength of associations among Web pages: (1) the hypertext [*sic*] link topology of a Web locality, (2) inter-page text similarity, and (3) the usage paths, or flow of users through the locality. Each of these networks or graphs is represented by matrices in our spreading activation algorithm.” (Pirolli, col. 10, lines 56-63.)

Further, Pirolli discloses creating and storing a combined graph that combines the graphs and that represents a generalized similarity of the documents inasmuch as Pirolli states that “[a]n

Art Unit: 2176

activation network can be represented as a graph defined by matrix R.” (Pirolli, col. 11, lines 36-37.)

Further, Pirolli discloses computing a generalized similarity value for a pair of documents based on the combined graph inasmuch as Pirolli states that “each off-diagonal element R_{ij} contains the strength of association between nodes i and j, and the diagonal contains zeros.” (Pirolli, col. 11, lines 36-39.)

Regarding **dependent claim 37**, Pirolli discloses classifying unclassified documents into category nodes of a taxonomy structure associated with the hypertext search system based on the generalized similarity value in combination with a comparison of a set of pre-classified training set of documents with a set of unclassified documents to carry out the classification inasmuch as Pirolli teaches selecting web pages for activation, the selected web pages being equivalent to training set documents (Pirolli, col. 5, lines 35-42 – “For relevancy predictions, one or more Web pages for spreading activation are selected, step 105. The selected Web pages may be based on the category that it is in. . . . In any event, activation is spread using the selected page as a focal point to generate a list of relevant pages, step 106.”), and further teaches use of the matrix R containing strength of association values, equivalent to similarity values for use in combination with the pre-classified documents. (Pirolli, col. 11, lines 35-59; *see also* Examples 1 and 2, col. 11, line 60 – col. 12, line 42.)

Claim Rejections - 35 USC § 103

13. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person

Art Unit: 2176

having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

14. **Claims 13-14, 17-18, and 27-32** are rejected under 35 U.S.C. 103(a) as being unpatentable over Pirolli.

Regarding **dependent claim 13**, Pirolli does not disclose measures of document similarity including URL similarity. However, Pirolli suggests using URL similarity as a measure of document similarity inasmuch as Pirolli teaches that the format and structure of documents' URLs as well as particular words found in documents' URLs might mean that the documents belong in the same category. (Pirolli, col. 9, lines 17-20, 24-28.) Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli to have used measures of document similarity including URL similarity.

Regarding **dependent claim 14**, Pirolli does not disclose considering two documents similar if a URL of each document contains similar URL sub-components. However, Pirolli suggests considering two documents similar if a URL of each document contains similar URL sub-components inasmuch as Pirolli teaches that particular words found in documents' URLs might mean that the documents belong in the same category. (Pirolli, col. 9, lines 24-28.) Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli to have considered two documents similar if a URL of each document contains similar URL sub-components.

Regarding **dependent claim 17**, Pirolli does not disclose achieving the combination of two or more measures of document similarity by taking the union of each of a plurality of graphs, each graph describing one of the measures of document similarity, to compute a combined graph that describes the combined document similarity. However, Pirolli suggests taking such a union

Art Unit: 2176

inasmuch as Pirolli states that “three kind of graphs, or networks, are used to represent strength of associations among Web pages: (1) the hypertext [*sic*] link topology of a Web locality, (2) inter-page text similarity, and (3) the usage paths, or flow of users through the locality. Each of these networks or graphs is represented by matrices in our spreading activation algorithm.” (Pirolli, col. 10, lines 56-63.) Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli to have taken such a union of a plurality of graphs.

Regarding **dependent claim 18**, Pirolli does not disclose achieving the combination of two or more measures of document similarity by taking the intersection of each of a plurality of graphs, each graph describing one of the measures of document similarity, to compute a combined graph that describes the combined document similarity. However, Pirolli suggests combining graphs inasmuch as Pirolli states that “three kind of graphs, or networks, are used to represent strength of associations among Web pages: (1) the hypertext [*sic*] link topology of a Web locality, (2) inter-page text similarity, and (3) the usage paths, or flow of users through the locality. Each of these networks or graphs is represented by matrices in our spreading activation algorithm.” (Pirolli, col. 10, lines 56-63.) Pirolli suggests that this combination could be an intersection inasmuch as Pirolli teaches that association strength can be zero, effectively meaning that a portion of a graph would be excluded from the combination. (Pirolli, col. 11, lines 1-34.) Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli to have taken such an intersection of a plurality of graphs.

Regarding **dependent claim 27**, Pirolli does not disclose categories derived from logs of user queries. However, Pirolli does suggest such a step inasmuch as Pirolli teaches that one of the three general sorts of information determine the need probabilities of information in memory,

Art Unit: 2176

given a current focus of attention [is] . . . past usage patterns” (Pirolli, col. 4, lines 34-35), and further explains that such usage patterns are found in “usage records or access logs of the web locality.” (Pirolli, col. 4, lines 58-59.) Therefore, it would have been obvious to one of ordinary skill in the art to have extended Pirolli to have derived categories from logs of user queries.

Regarding **dependent claim 28**, Pirolli does not disclose creating and storing the matrix using columns representing documents and rows representing user sessions wherein values of elements of the second matrix represent interest in a document shown by a particular user in a particular session. However, Pirolli does teach creating and storing information about the number of times a document was requested within a given time period (Pirolli, col. 8, lines 24-25) and also suggests tracking user usage patterns (Pirolli, col. 4, lines 34-35), which suggests creating and storing a value that is the function of the amount of time a user spent viewing a document associated with a particular session. Therefore, it would have been obvious to one of ordinary skill in the art to have extended Pirolli to have created columns representing documents and rows representing user sessions wherein values represent interest in a document shown by a particular user in a particular session.

Regarding **dependent claim 29**, Pirolli does not disclose creating and storing the matrix using rows representing documents and columns representing user sessions wherein values of elements of the second matrix represent interest in a document shown by a particular user in a particular session. However, Pirolli does teach creating and storing information about the number of times a document was requested within a given time period (Pirolli, col. 8, lines 24-25) and also suggests tracking user usage patterns (Pirolli, col. 4, lines 34-35), which suggests creating and storing a value that is the function of the amount of time a user spent viewing a

Art Unit: 2176

document associated with a particular session. Therefore, it would have been obvious to one of ordinary skill in the art to have extended Pirolli to have created rows representing documents and columns representing user sessions wherein values represent interest in a document shown by a particular user in a particular session.

Regarding **dependent claim 30**, Pirolli does not disclose element values computed as a function of a time that a user has spent viewing a document associated with each element. However, Pirolli does teach creating and storing information about the number of times a document was requested within a given time period (Pirolli, col. 8, lines 24-25) and also suggests tracking user usage patterns (Pirolli, col. 4, lines 34-35), which suggests creating and storing a value that is the function of the amount of time a user spent viewing a document associated with a particular session. Therefore, it would have been obvious to one of ordinary skill in the art to have extended Pirolli to have computed element values as a function of a time that a user has spent viewing a document associated with each element.

Regarding **dependent claim 31**, Pirolli discloses creating and storing a second matrix representing a Similarity between pairs of documents i and j wherein the second matrix is derived by comparing pairs of column vectors or row vectors respectively i and j of the first matrix inasmuch as Pirolli teaches generating three matrices representing similarity between documents (Pirolli, col. 10, lines 10-11; Figs. 9, 11, 13) from raw information entered in a first matrix (Pirolli, Fig. 5).

Regarding **dependent claim 32**, Pirolli discloses creating and storing a second matrix representing a Similarity between pairs of documents i and j inasmuch as Pirolli teaches generating three matrices representing similarity between documents (Pirolli, col. 10, lines 10-

Art Unit: 2176

11; Figs. 9, 11, 13) from raw information entered in a first matrix (Pirolli, Fig. 5). Pirolli does not disclose finding pairs of documents *i* and *j* which have high interest values for a particular user in a particular session or period of time. However, Pirolli does teach creating and storing information about the number of times a document was requested within a given time period (Pirolli, col. 8, lines 24-25) and also suggests tracking user usage patterns (Pirolli, col. 4, lines 34-35), which suggests comparing documents based on interest values for a particular user in a particular session of time. Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli to have created the second matrix as recited in claim 32.

15. **Claims 15-16** are rejected under 35 U.S.C. 103(a) as being unpatentable over Pirolli in view of U.S. Patent Number 6,282,549 B1 to Hoffert et al. (hereinafter "Hoffert"), issued August 28, 2001, filed March 29, 1999.

Regarding **dependent claim 15**, Pirolli does not disclose measures of similarity including multimedia similarity. Hoffert, however, teaches that the type of multimedia file is relevant to a user querying a database for multimedia files, and teaches the classification of multimedia files with associated icons indicating file type. (Hoffert, col. 23, lines 46-67.) Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli to have measures of similarity include multimedia similarity.

Regarding **dependent claim 16**, Pirolli does not disclose considering two documents similar based on features derived from multimedia components linked to or contained by the documents. Hoffert, however, teaches the storage of a variety of multimedia file features for the storage, retrieval, and classification of multimedia documents. (Hoffert, col. 6, lines 10-32.) Therefore, it would have been obvious to one of ordinary skill in the art to have modified Pirolli

Art Unit: 2176

to consider two documents similar based on features derived from multimedia components linked to or contained by the documents.

16. **Claim 20** is rejected under 35 U.S.C. 103(a) as being unpatentable over Pirolli in view of U.S. Patent Number 6,128,606 to Bengio et al. (hereinafter “Bengio”), issued October 3, 2000, filed March 11, 1997.

Regarding **dependent claim 20**, Pirolli does not disclose obtaining structural information by optimizing an objective function. However, Bengio, in disclosing an invention “directed to the problem of developing a modular building block for complex processes that can input and output data in a wide variety of forms, but when interconnected with other similar modular building blocks can be easily trained” (Bengio, col. 2, lines 45-49), teaches “training a network of these modules by back-propagating gradients through the network to determine a minimum of the global objective function.” (Bengio, col. 2, lines 57-60.) Because claim 20 is directed to a similar invention, it would have been obvious to one of ordinary skill in the art to have combined Pirolli and Bengio to implement the optimization of an objective function.

17. **Claims 21-25** are rejected under 35 U.S.C. 103(a) as being unpatentable over Pirolli in view of U.S. Patent Number 6,389,436 to Chakrabarti et al. (hereinafter “Chakrabarti”), issued May 14, 2002, filed December 15, 1997.

Regarding **dependent claim 21**, Pirolli does not disclose obtaining structural information by approximately optimizing an objective function. Chakrabarti, however, in the context of a document classifier similar to the invention of claim 21, discloses optimizing an objective function by “relaxation labeling” in which “[t]he iteration continues until a stopping criteria is reached.” (Chakrabarti, col. 19, lines 17-20.) Therefore, it would have been obvious to one of

Art Unit: 2176

ordinary skill in the art to have combined Pirolli and Chakrabarti to have obtained structural information by approximately optimizing an objective function.

Regarding **dependent claim 22**, neither Pirolli nor Chakrabarti discloses repeated application of a growth transformation. However, given that a growth function is one which by definition stabilizes in a finite number of steps, it would have been obvious for one of ordinary skill in the art to have extended the combination of Pirolli and Chakrabarti to repeatedly apply a growth transformation.

Regarding **dependent claim 23**, Pirolli does not disclose creating and storing a second matrix that represents an interim score for each document in each category. However, Chakrabarti teaches a technique of soft classification in which “after each iteration, all documents are assigned a vector containing estimated probabilities of belonging to each class.” (Chakrabarti, col. 19, lines 27-29.) Therefore, it would have been obvious to one of ordinary skill in the art to have combined Pirolli and Chakrabarti to have created and stored a second matrix that represents an interim score for each document in each category.

Regarding **dependent claim 24**, Pirolli does not disclose periodically normalizing the rows of the matrix by normalizing within each document, across all categories, whereby the score for one document in a particular category will depend on the scores for that document in other categories. However, Chakrabarti suggests such a step inasmuch as Chakrabarti teaches word vectors containing probabilities in which the score for one document in a particular category inherently depends on the scores for that document in all other categories. (Chakrabarti, col. 6, lines 46-65; col. 19, lines 27-29.) Therefore, it would have been obvious to one of

Art Unit: 2176

ordinary skill in the art to have combined Pirolli and Chakrabarti to have periodically normalized the rows of the matrix as recited in claim 24.

Regarding **dependent claim 25**, Pirolli does not disclose periodically, as the matrix is being computed, normalizing columns of the matrix by normalizing within each category, across all documents, whereby the score for one document in a particular category depends on the scores for all other documents in that category. However, Chakrabarti suggests such a step inasmuch as Chakrabarti teaches word vectors containing probabilities in which the score for one document in a particular category inherently depends on the scores for that document in all other categories. (Chakrabarti, col. 6, lines 46-65; col. 19, lines 27-29.) Therefore, it would have been obvious to one of ordinary skill in the art to have combined Pirolli and Chakrabarti to have periodically normalized the columns of the matrix as recited in claim 25.

Conclusion

18. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

United States Patents:

6,446,099 B1	6,418,432 B1
6,418,431, B1	6,405,188 B1
6,397,212 B1	6,360,215 B1
6,338,060 B1	6,321,220 B1
6,282,549 B1	6,182,091 B1
6,182,066 B1	6,115,718
6,112,203	6,052,676

Art Unit: 2176

5,960,422

5,931,907

5,920,859

5,895,470

5,799,292

5,745,893

Other documents:

Hermann Kaindl, Stefan Kramer, Luis Miguel Afonso, *Combining Structure Search and Content Search for the World-Wide Web*, Proceedings of the Seventh ACM Conference on Hypertext and Hypermedia, pp. 217-224 (1998).

Ron Weiss, Bienvenido Vélez, Mark A. Sheldon, *HyPursuit: A Hierarchical Network Search Engine That Exploits Content-Link Hypertext Clustering*, Proceedings of the Seventh ACM Conference on Hypertext, pp. 180-193 (1996).

19. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Charles A. Bieneman whose telephone number is 703-305-8045. The examiner can normally be reached on Monday - Thursday, 7:00 a.m. - 5:30 p.m..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 703-308-5186. The fax phone numbers for the organization where this application or proceeding is assigned are 703-746-7239 for regular communications and 703-746-7238 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is 703-305-4700.

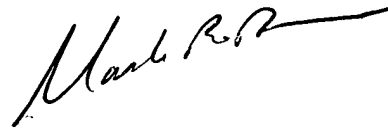
Application/Control Number: 09/333,121

Page 22

Art Unit: 2176

CAB

October 23, 2002

A handwritten signature in black ink, appearing to read "Mark R. Powell". The signature is written in a cursive style with a long horizontal stroke at the end.

MARK R. POWELL
SUPERVISORY PATENT EXAMINER
ART UNIT 2415